Research Article

Numerical Computation of pH and Buffer Capacity in Complex Mixtures of Acids, Bases and Ampholytes

Wibul Wongpoowarak, Damrongsak Faroongsarng, Nimit Worakul and Prapaporn Boonme*

Department of Pharmaceutical Technology, Faculty of Pharmaceutical Sciences,
Prince of Songkla University, Hat-Yai, Songkhla, Thailand
**Corresponding author. E-mail address: prapaporn.b@psu.ac.th*

Received September 14, 2011; Accepted May 5, 2012

Abstract

This study was aimed to simplify the pH calculation for complex mixtures of acids, bases and ampholytes; the theoretical concept on pH calculation was renovated. A new algorithm was proposed to state that buffer capacity (β) could be simultaneously achieved with pH. To validate the method, the buffer systems of Na_2HPO_4/NaH_2PO_4 and Na_2HPO_4/citric acid were employed. Acid and base titration was done on the buffer systems using HCl and NaOH, respectively. It was demonstrated that only the charge balance equation (CBE) was necessary and sufficient to successfully compute β. To achieve the desirable accuracy, the pH interval used in such algorithm should be 10^{-4} for general calculation; between 10^{-4} and 10^{-2} for 8 significant-digits calculation; between 2×10^{-10} and 4×10^{-4} for 16 significant-digits calculation. Visual Basic Source Code written in Microsoft Access™ for demonstrating computation of pH and β of pure water was also illustrated.

Key Words: pH; Buffer capacity; Computational prediction; Algorithm

Introduction

Prediction of pH and buffer capacity (β) of multi-component solutions has been very useful in several applications. For pharmaceutical field, it can be used in the design of dosage forms (York, 2002), in analytical process (Heyrman and Henry, nd), in prediction of oral drug absorption (Obata et al., 2005) and in food preservation (Coiffard et al., 1997; Chung et al., 2010; Garcia-Gonzalez et al., 2010).

A method in pH and β calculation is algebraic method under specific assumptions for each system. This assumption-dependent algebraic method is inconvenient to predict pH of a multi-component solution. Butler (1964) described a unified concept to calculate pH of a solution by writing and solving two equations from three possible equations, i.e., proton balance equation (PBE), mass balance equation (MBE) and charge balance equation (CBE). The variables appeared in these equations relate to each other by dissociation constant of solutes in a solvent. If these relationships could be solved and as long as no constraints were involved, pH prediction of a solution containing various acids, bases and/or ampholytes would be possible. If in the situation with constraints, other parameters should be mutually considered. Nevertheless, solving simultaneous equations would be ending up with high-degree nonlinear-multivariate polynomial equations. Since the root of the equation must be solved with numerical method, it may not be suitable for database purpose due to the complexities in the calculation process.

The database could be constructed on the basis that we could numerically and symbolically split substance dissolved into all related species. By varying pH, the pK_a, pK_w and mass balance relationship could be solved for all species at that particular pH. By using the approach suggested by Cutler (1986), pH was guessed in binary search style to check for validity of charge balance equation. When the valid pH was found (no charge balance anomaly), buffer capacity is also know simultaneously according to our algorithm.

The aim of the article was to demonstrate that not only β could be directly derived from CBE with slight modification of Cutler's concept but also simultaneously obtained with pH in the prediction process.

Review of the computation of pH and β

Theoretical aspects of CBE

In multi-component mixtures, it was assumed that there was no interference between different species via chemical reactions. At any specified pH, all species conforming to that pH could be able to compute from the definition of weak acid dissociation constants (K_a's), dissociation constant of water (K_w) and MBE. In the other words, the pH could be solved from MBE, CBE and dissociation constants (Butler, 1964). According to an electroneutrality rule, the valid pH would be where the positive-charge side equaled to the negative-charge side in the CBE (Cutler, 1986). The difference of both sides calculated from the negative-charge side minus the positive-charge side in CBE, designated as Δ, is then introduced as an indicator for the validity of CBE. The CBE is valid when the value of Δ is zero and that pH would be the unique possible value of such a system. Moreover, with pH increases, negative charge increases whereas positive charge decreases. The value of Δ should be a monotonous function of pH. This monotonous property is specifically solvable by binary search method (Cutler, 1986).

Since the charge balance is always defined in terms of equivalent unit, the negative and positive species are referred to basic and acidic compounds, respectively. Therefore, Δ is the base in excess (i.e., equivalent to adding bases into a system). The slope of the plot between the pH and the corresponding Δ (where pH exhibits the actual value, i.e., at Δ = 0) is the change of bases in terms of equivalent unit per liter versus the change in pH from its initial equilibrium condition. Thus, β will be the slope of the Δ vs pH plot *by definition*, since β is defined as the ratio of change of bases in equivalent unit per liter divided by the change in pH. By finding the actual pH according to Cutler's concept, the β of that system could also be obtained simultaneously from the slope of Δ-pH plot at that specific pH.

Since Δ could be interpreted as the base in excess that corresponds to the changing of pH, the reverse was also expected to be true, i.e., the base in excess caused the pH to change. It implied that the whole isovolumetric titration profile of any fixed system could be constructed within the single-point pH calculation process.

Species calculation

At any given pH, all related species could be computed from algebraic expressions (Sinko, 2010). For a complex mixture of acids, bases and ampholytes, the expressions may be complicated.

It is of importance to have an insight into the calculation process for triprotic acid with the simplest algebraic expression that is more suitable to implement into software. The species calculation is arranged into 2 steps.

In the first step, an arbitrary unit could be used in the process, e.g., H_3A = 1 unit. The next species will be the previous species multiply with 10^{pH-pKa}. For the particular example, all species involved would be as illustrated in Eq. 1-3, where H_3A is a representative of an acid with three hydrogen atoms which can be dissociated to H_2A^-, HA^{2-}, and A^{3-} (Sinko, 2010).

$$H_2A^- = H_3A \times 10^{pH-pK_{a1}} \quad \text{(Eq. 1)}$$

$$HA^{2-} = H_2A^- \times 10^{pH-pK_{a2}} \quad \text{(Eq. 2)}$$

$$A^{3-} = HA^{2-} \times 10^{pH-pK_{a3}} \quad \text{(Eq. 3)}$$

The overall sum of all HA-related species (T) can be calculated using Eq. 4.

$$T = H_3A + H_2A^- + HA^{2-} + A^{3-} \quad \text{(Eq. 4)}$$

In the second step, correction with the actual unit was done. Let M be the total concentration of all HA-related species,

$$T \text{ unit} = M \text{ molar} \quad \text{(Eq. 5)}$$

By multiplying with M/T, all species in molar unit could be obtained. In computational software program, at an assigned pH, species calculation is straightforward even in a solution containing multicomponents or polyelectrolytes, using only the basic definition of the acidity constants involved.

Classical β calculation

In the case of weak acids or weak bases in aqueous solution, the exact Van Slyke's equation showed in Eq. 6 is utilized (Sinko, 2010):

$$\beta = \ln(10)\{[H_3O^+] + [OH^-] + \frac{C[H_3O^+]K_a}{([H_3O^+] + K_a)^2}\} \quad \text{(Eq. 6)}$$

where β and C are buffer capacity and the total concentration of that buffer species, respectively. In the case of overlapping pK_a near the pH of interest, buffer capacity contribution from all overlapping pK_a had to be summed up together, i.e., insertion of summation sign for the last terms is a more appropriate form for such case.

Theoretically, the contribution of pH on the β of the system is determined according to Eq. 7, i.e., ignoring the last term in the Eq. 6.

$$\beta = \ln(10)([H_3O^+] + [OH^-]) \quad \text{(Eq. 7)}$$

The Eq. 7 is pH effect, which could be totally neglected at physiological pH. The minimum β would be 4.6 x 10^{-7} at pH 7 (see Appendix for demonstrating code to compute this value by using the principle outlined in the above section). At physiological pH, suitable for pharmaceutical preparation where pH is not in too acidic or basic region, only the third term in the Eq. 6 can be expanded by using the definition of pH and pK_a to yield Eq. 8.

$$\beta = \ln(10)\frac{C(10^{-pH-pK_a})}{(10^{-2pH} + 2\times 10^{-pH-pK_a} + 10^{-2pK_a})} \quad \text{(Eq. 8)}$$

Simplifying the Eq.8 by factoring with $(10^{-pH-pKa})$, the Eq. 9 is gotten.

$$\beta = \ln(10)\frac{C}{(10^{pK_a-pH} + 2 + 10^{pH-pK_a})} \quad \text{(Eq. 9)}$$

Rearranging the terms in the parenthesis as well as substituting the term $10^{|pH-pKa|}$ by *X*, a simplified form of Van Slyke's equation for β (without pH effect), Eq. 9 can be rewritten as Eq. 10.

$$\beta = \ln(10)(\frac{C}{2 + X + \frac{1}{X}}) \quad \text{(Eq.10)}$$

In a mixture of various buffering components, the total β could be computed from its partial contribution for each dissociation constant. In addition, the β correction due to the pH effect (Eq. 7), together with the ionic strength correction, may be used to achieve more accurate value.

Materials and Methods

Materials

Sodium dihydrogen phosphate monohydrate ($NaH_2PO_4.H_2O$), hydrochloric acid (HCl) and sodium hydroxide (NaOH) were obtained from Merck, Germany. Di-sodium hydrogen orthophosphate (Na_2HPO_4) were obtained from BDH Laboratory Supplied, UK. Citric acid ($C_6H_8O_7$) was obtained from Vidhyasom Co., Ltd., Thailand. Distilled water was used throughout the experiment. All chemicals were pharmaceutical grade and used without further modification.

Determination of pH and β of solution containing 0.002 M Na_2HPO_4 and 0.001 M NaH_2PO_4

This experiment was done by the titration of buffer solution with acid or base. The acidic or basic buffer solutions were prepared by adding a stock solution of HCl or NaOH previously standardized following the method specified in USP24 (2000) into buffer solution to get a final concentration of 0.02 gram-equivalent/liter. The concentrations of Na_2HPO_4 and NaH_2PO_4 in the solutions were similar to those of original buffer, i.e., 0.002 M Na_2HPO_4 and 0.001 M NaH_2PO_4. During titration of these solutions into an original buffer, the concentration of HCl or NaOH gradually changed whereas the sum of Na_2HPO_4 and NaH_2PO_4 concentrations was kept constant. The titrant was added by micropipette and stirred with magnetic stirrer (Model MR 3000D, Heidolph, Germany). The sample was allowed to equilibrate before each pH recording by a pH meter (Model 20, Denver Instrument Company, USA). Titration was performed at 25°C in three replicates. The concentration of acid (negative value) or base (positive value) added were plotted against corresponding pH. The β was calculated from the slope of a linear regression when small amount of acid or base added. These results were compared to the calculated pH and β values used the binary search method with optimizing computation accuracy.

Determination of pH and β of solution containing X ml of 0.2 M Na_2HPO_4 and (100-X ml) of 0.1 M citric acid

Buffer solutions with different volume fractions of 0.2 M Na_2HPO_4 and 0.1 M citric acid were prepared and determined pH values at 25°C using the pH meter. The β was determined by the identical method as previously mentioned. The experimental values were plotted with the simulated values and the values taken from the reference (Dean, 1999) to demonstrate the applicability of the proposed method in prediction of the pH and β of a complicated mixture, in which the simultaneous equations may not be easily solved.

Results and Discussion

Calculation for β: optimizing computation accuracy

As stated by Culter (1986), MBE and CBE as multivariate equations could be transformed to a univariate equation. However, solving univariate problem by numerical method is not always flexible for software implementation, some method may also posses a limitation, for instance, Newton-Raphson method may encounter numerical instability in a very low slope region (Mathews, 1987). Since low slope would occur in the case of very low β system, Newton-Raphson method would not be a preferable choice. The binary search method was selected since its stable with monotonous function and always rapidly converged to 8 significant-digits precision within 30 iterations.

Since the slope is a derivative of a function, approximation of slope requires infinitely small value of pH interval. However, too small value of this pH interval may lead to numerical rounding error as a trade-off. In the binary search, the interval for each of iteration could be reduced by half. When the search interval was near an optimal value, one would achieve the highest calculation accuracy. To determine its effect on the accuracy of β calculation, the computer simulation was performed using protocol shown in Algorithm 1.

Algorithm 1 Simulation protocol for finding optimal pH interval to be used in β computation process.

Vary acid concentration within 1 and 10^{-15} range.
Vary pH interval within 1 and 10^{-11} range.
Calculate pH of the system using such pH interval
Calculate β by this method using such pH interval
Calculate β by exact Van Slyke's equation
Calculate prediction error
Record the best pH interval at each acid concentration
Summarize the optimal pH interval that is applicable to most cases of acid concentrations.
Note: Criteria to determine the best pH interval is to minimize MaxError or RMSE

$$\mathrm{MaxError} = \mathrm{Maimum}\ (\left|\beta_{\mathrm{ThisMethod}} - \beta_{\mathrm{VanSlyke}}\right|)$$

$$\mathrm{RMSE} = \sqrt{(\sum \beta_{\mathrm{ThisMethod}} - \beta_{\mathrm{VanSlyke}})^2 / \mathrm{N}}$$

Optimal pH interval is the pH interval that cause the calculation error to be smaller than the acceptable limit:
MaxError $\leq$ Multiplier x best case MaxError and RMSE $\leq$ Multiplier x best case RMSE
MaxError is the worst-case prediction error, while RMSE is the general prediction error in statistical sense.
The multiplier of 10 is used in nearly all conditions, except when the optimal pH interval range is too narrow for practical purpose or shift away from the usual range applicable to other situations that we will use the multiplier of 100 instead.

In addition, the optimal pH intervals for determining the β resulting from each protocol are shown in Algorithm 2.

Algorithm 2 Optimizing results using the protocol according to Algorithm 1.

1. Strong monoprotic acid at various concentrations between 1 and 10^{-15}
   - Single precision computation mode
     - *Conclusion:*
       By using the multiplier of 10 (see Algorithm 1), the optimal pH range is very narrow and may limit its usefulness; therefore multiplier as 100 is used to obtain a more applicable range. It was found that the optimal pH interval should be between 3 x 10^{-5} and 5 x 10^{-2}.
   - Double precision computation mode
     - *Conclusion:*
       The optimal pH interval should be between 2 x 10^{-10} and 4 x 10^{-4}.
2. Weak monoprotic acid at various concentration between 1 and 10^{-15}
   - Single precision computation mode
     - *Conclusion:*
       The optimal pH interval should be between 1 x 10^{-4} and 1 x 10^{-2}.
   - Double precision computation mode
     - *Conclusion:*
       The optimal pH interval should be between 2 x 10^{-10} and 4 x 10^{-4}.

Conclusion: The universal optimal pH interval is 1 x 10^{-4}. This should be applicable for both single and double precision calculation. In the case of calculation with specific significant digit, this optimal value should be changed.

We found that pH interval between 10^{-4} and 10^{-2} was suitable for single precision mode (8 significant digits), and that between 2 x 10^{-10} and 4 x 10^{-4} for double precision mode (16 significant digits). The value of 10^{-4} was set since it could be used satisfactorily for both single and double precision calculation in various situations. The computed β is accurate for at least 3 decimal places. If the selected pH interval was higher than this optimal value, the accuracy may be deteriorated, whereas, smaller value would lead to rounding error instability. In addition, the method allowed computing β simultaneously with the pH value. Algorithm 3 summarized a pseudo-code for program implementation.

pH and β of a solution containing 0.002 M Na_2HPO_4 and 0.001 M NaH_2PO_4

Approximating the pH of this system using Henderson-Hasselbalch equation, by assuming the pH close to pK_{a2} due to the fact that $[HPO_4^{2-}]$ and $[H_2PO_4^-]$ are the dominant species, provides pH at 7.51 (= pK_{a2} + log 0.002/0.001). The β of this system computed from Van Slyke's equation with correction from pH effect was 0.00153.

A group of equations containing those of dissociation constants, MBE, and CBE was specified according to Algorithm 4. By varying pH and computing all related species, a plot of pH against Δ in Figure 1 could be illustrated. It was found that exact solving for the root of equation (pH at which Δ = 0) provided a pH of 7.38±0.02, with slope at that point equaled to 0.00151±0.00004 (i.e., the β of the system). These values were comparable with those obtained from Van Slyke's equation.

Algorithm 3 The pseudo-code of buffer computation for program implementation.

```
Set Maximum pH and compute CBE inequality at Maximum pH (usually 14 or higher).
Set Minimum pH and compute CBE inequality at Minimum pH (usually 0 or lower).
Set Converge = False
Do
    Set Middle pH to (Maximum pH + Minimum pH)/2
    Compute CBE inequality at Middle pH
    If CBE inequality at Middle pH is positive then
        Set Maximum pH to Middle pH (and update its CBE inequality)
    Else
        Set Minimum pH to Middle pH  (and update its CBE inequality)
    End if
    If Maximum pH - Minimum pH > 0.0001 (the optimal pH interval) then
        Set $\beta$ to (CBE inequality at Maximum pH - CBE at Minimum pH) divided by (Maximum pH - Minimum pH)
    Else
        Set Converge = TRUE
    End if
Loop until Converge
Report Middle pH and the corresponding $\beta$
```

Algorithm 4 Equations dealing with dissociation constants as well as MBE and CBE of 0.002 M Na_2HPO_4 / 0.001 M NaH_2PO_4 buffer.

Constant definition: $K_w = [H_3O^+][OH^-]$

$$K_{a1} = [H_3O^+][H_2PO_4^-]/[H_3PO_4] = 10^{-2.12}$$

$$K_{a2} = [H_3O^+][HPO_4^{2-}]/[H_2PO_4^-] = 10^{-7.21}$$

$$K_{a3} = [H_3O^+][PO_4^{3-}]/[HPO_4^{2-}] = 10^{-12.67}$$

Mass balance: $[H_3PO_4] + [H_2PO_4^-] + [HPO_4^{2-}] + [PO_4^{3-}] = 0.003$

$$[Na^+] = 0.005$$

Charge balance: $[H_3O^+] + [Na^+] = [OH^-] + [H_2PO_4^-] + 2[HPO_4^{2-}] + 3[PO_4^{3-}]$

From charge balance, Δ (the deviation from the electroneutrality rule) is described as the negative charge side of CBE minus the positive-charge side of CBE.

$$\Delta = ([OH^-] + [H_2PO_4^-] + 2[HPO_4^{2-}] + 3[PO_4^{3-}]) - ([H_3O^+] + [Na^+])$$

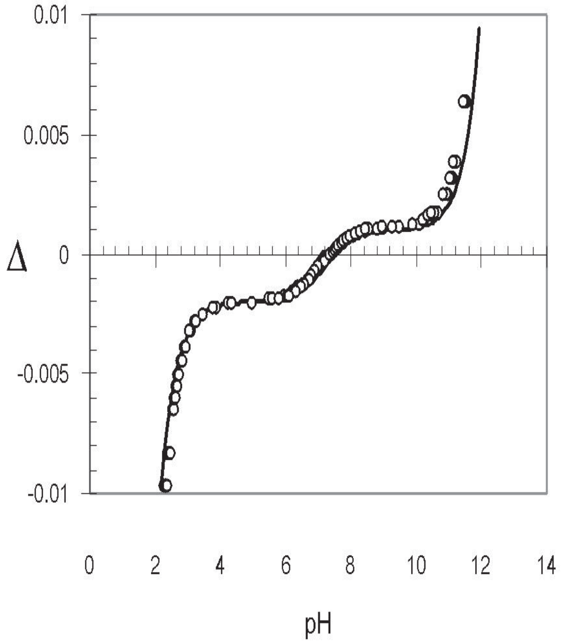

Figure 1 Plot of pH against Δ of 0.002 M Na_2HPO_4/0.001 M NaH_2PO_4 buffer from computer simulation using conditions in Algorithm 4: circular symbol referred to experimental values and solid lines referred to calculated values

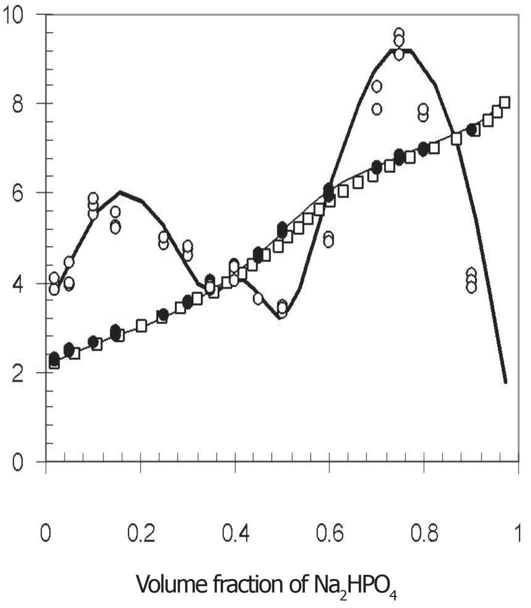

Figure 2 Plot between volume of 0.2 M Na_2HPO_4 against pH and β of Na_2HPO_4/citric acid buffer: circular symbol referred to experimental values (blank: 100β, filled: pH), square symbol referred to pH values taken from the reference (Dean, 1999), and solid lines referred to calculated values

pH and β of a solution containing x mL of 0.2 M Na_2HPO_4 plus (100-x) mL of 0.1 M Citric acid

A set of pH's and β's of a solution containing various volume fraction of 0.2 M Na_2HPO_4 and that of 0.1 M citric acid was computed using the previous manner. Figure 2 illustrates the plot of volume fraction of 0.2 M Na_2HPO_4 against pH and 100β for the ease of viewing.

In addition to the computation, Robinson-Guggenheim-Bates correction for pH due to ionic strength (Mathews, 1987) was utilized. It was observed that the experimental data of the calculated pH and β were comparable to those of the available pH data previously published (Dean, 1999). It was clearly demonstrated that the overall concept was satisfactorily applicable. The pH and β of a solution containing mixtures of polyprotic acids that were generally regarded as difficult and very tedious to calculate could be easily accessible by this method.

Conclusions

The concept of Butler and Cutler would be applicable to predict pH of a buffer without any difficulty for a multicomponent system. However, when precipitation or chemical change occurs, the usefulness of this concept is limited due to the effects of the mass balance. The precipitation or chemical change will be the constraint of solving mass balance equation. The binary search method would become useless in such case. Multivariable search methods should be used instead, which is outside the scope of this paper.

In this article, the charge balance equation contained necessary and sufficient information to compute β was proposed. With proper implementation, the accuracy of computation was ascertained. From the simplicity of these concepts, it was possible to create a database that could compute various solution properties relating to pH. When systematically combined with pK_a prediction by group contribution methods, it is possible to make a priori prediction of various solution properties that is necessary for separation in combinatorial chemistry. The database could be customized to be used in experimental research such as customizing buffer with desirable properties.

References

Butler, J. N. (1964) *Ionic Equilibrium: A Mathematical Approach*, Addison-Wesley, Massachusetts, pp.61-68.

Chung, C. C., Chen, H. H., and Ting, C. H. (2010) Grey prediction fuzzy control for pH processes in the food industry. *Journal of Food Engineering* 96: 575-582.

Coiffard, C. A., Coiffard, L. J., and De Roeck-Holtzhauer, Y. M. (1997) Influence of pH on thermodegradation of thaumatin in aqueous solution. *Food Research International* 30: 707-710.

Cutler, D. (1986) Calculation of pH for complex mixtures of acids, bases and ampholytes. *Journal of Pharmacy and Pharmacology* 38: 499-501.

Dean, J. A. (1999) *Lange's Handbook of Chemistry*, 15th ed., McGraw-Hill, New York.

Garcia-Gonzalez, L., Teichert, H., Geeraerd, A. H., Elst, K., Van Ginneken, L., and Van Impe, J. F., Vogel, R. F., Devlieghere, F. (2010) Mathematical modelling and in situ determination of pH in complex aqueous solutions during high-pressure carbon dioxide treatment. *The Journal of Supercritical Fluids* 55: 77-85.

Heyrman, A. N. and Henry, R. A. (nd), *Importance of controlling mobile phase pH in reversed phase HPLC*. [Online URL: www. hplcsupply. com/pdf/App_9.pdf.] accessed on January 14, 2011.

Mathews, J. M. (1987) *Numerical Methods for Computer Science, Engineering, and Mathematics*, pp.64-65. Prentice-Hall, London.

Obata, K., Sugano, K., Saitoh, R., Higashida, A., Nabuchi Y., Machida, M., and Aso, Y. (2005), Prediction of oral drug absorption in humans by theoretical passive absorption model. *International Journal of Pharmaceutics* 293: 183-192.

Sinko, P. J. (2010) *Martin's Physical Pharmacy and Pharmaceutical Sciences*, 6th ed., pp. 163-181. Lippincott Williams & Wilkins, Philadelphia.

USP24 (2000) *The United States Pharmacopeia 24*, pp.2231-2232. The United States Pharmacopeial Convention, Maryland.

York, P. (2002) The design of dosage forms, In *Pharmaceutics: The Science of Dosage Form Design* (Aulton, M.E., ed.), 2nd ed., pp.1-12. Churchill Livingstone, London.

Appendix

Demonstration of computing pH and β of water

Constant definition: $K_w = [H_3O^+][OH^-]$

Mass balance: $[H_3O^+] = 10^{-pH}$

$$[OH^-] = K_w / [H_3O^+]$$

Charge balance: $[H_3O^+] = [OH^-]$

From charge balance, D (the deviation from the electroneutrality rule) is described as the negative charge side of CBE minus the positive-charge side of CBE.

$$\Delta = ([OH^-] - [H_3O^+])$$

The following code was written in the Microsoft Access™. Please be noted that log ( ) function is natural logarithm.

```
Option compare database
DefDbl A-Z
Sub Main()
    minph = 0
    maxph = 14
    Kw = 1e-14
    Do
        pH = (maxph + minph) / 2
        H = 1 / 10 ^pH
        OH = Kw / H
        Delta = OH - H
        If Delta > 0 Then
          maxph = pH
          MaxpHDelta = Delta
         Else
          minph = pH
          MinpHDelta = Delta
        End If
        If maxph - minph > 0.00001 Then
            ‘Remark: optimal pH interval is 0.00001 for double precision mode
             Beta = (MaxpHDelta - MinpHDelta) / (maxph - minph)
        Else
           Converge = True
           Exit Do
        End If
    Loop Until Converge
    Debug.Print “pH =”; pH
    Debug.Print “β (our method):"; Beta
     Debug.Print "β (Van Slyke):"; Log(10) * (H + OH)
    End Sub
```

By typing **main** and press [Enter] key (in the Debug window), the following result would be shown in the Debug Window.

```
pH = 6.99999332427979
β (our method): 4.6051701867266E-07
β (Van Slyke): 4.60517018653215E-07
```